

Naive Bayes

Nipun Batra

IIT Gandhinagar

July 29, 2025

$$P(x_1, x_2, x_3, \dots, x_N | y) = P(x_1 | y) P(x_2 | y) \dots P(x_N | y)$$

$$P(x_1, x_2, x_3, \dots, x_N | y) = P(x_1 | y) P(x_2 | y) \dots P(x_N | y)$$

Why is Naive Bayes model called Naive?

$$P(x_1, x_2, x_3, \dots, x_N | y) = P(x_1 | y) P(x_2 | y) \dots P(x_N | y)$$

Why is Naive Bayes model called Naive?

Naive assumption x_i and x_{i+1} are independent given y

$$\text{i.e. } p(x_2 | x_1, y) = p(x_2 | y)$$

$$\begin{aligned}
 & P(y = 1 | w_1 = 0, w_2 = 0, w_3 = 1) \\
 = & \frac{P(w_1 = 0 | y = 1)P(w_2 = 0 | y = 1)P(w_3 = 1 | y = 1)P(y = 1)}{P(w_1 = 0, w_2 = 0, w_3 = 1)} \\
 = & \frac{0.6 \times 0.8 \times 0.6 \times 0.5}{Z}
 \end{aligned}$$

$$\begin{aligned}
 & P(y = 1 | w_1 = 0, w_2 = 0, w_3 = 1) \\
 = & \frac{P(w_1 = 0 | y = 1)P(w_2 = 0 | y = 1)P(w_3 = 1 | y = 1)P(y = 1)}{P(w_1 = 0, w_2 = 0, w_3 = 1)} \\
 = & \frac{0.6 \times 0.8 \times 0.6 \times 0.5}{Z}
 \end{aligned}$$

Similarly, we can calculate

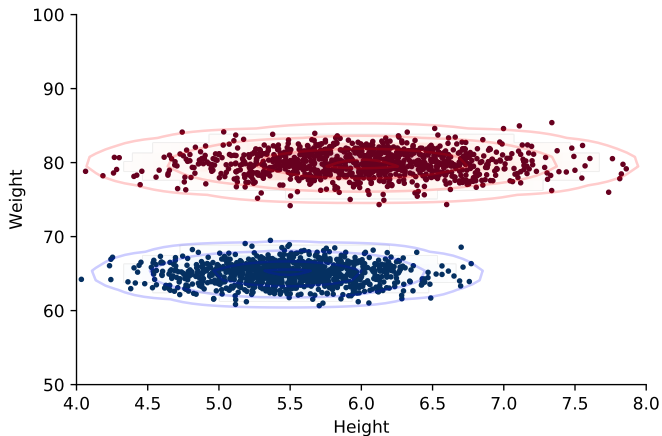
$$P(y = 0 | w_1 = 0, w_2 = 0, w_3 = 1) = \frac{0.6 * 0.4 * 0.6 * 0.5}{Z}$$

$$\begin{aligned}
 & P(y = 1 | w_1 = 0, w_2 = 0, w_3 = 1) \\
 = & \frac{P(w_1 = 0 | y = 1)P(w_2 = 0 | y = 1)P(w_3 = 1 | y = 1)P(y = 1)}{P(w_1 = 0, w_2 = 0, w_3 = 1)} \\
 = & \frac{0.6 \times 0.8 \times 0.6 \times 0.5}{Z}
 \end{aligned}$$

Similarly, we can calculate

$$\begin{aligned}
 P(y = 0 | w_1 = 0, w_2 = 0, w_3 = 1) &= \frac{0.6 \times 0.4 \times 0.6 \times 0.5}{Z} \\
 \frac{P(y=1|w_1=0,w_2=0,w_3=1)}{P(y=0|w_1=0,w_2=0,w_3=1)} &= 2 > 1. \text{ Thus, classified as a spam} \\
 &\text{example.}
 \end{aligned}$$

Note: no cross covariance! Remember all features are independent.



Wikipedia Example

Height	Weight	Footsize	Gender
6	180	12	M
5.92	190	11	M
5.58	170	12	M
5.92	165	10	M
5	100	6	F
5.5	100	6	F
5.42	130	7	F
5.75	150	7	F

Example

	Male	Female
Mean (height)	5.855	5.41
Variance (height)	3.5×10^{-2}	9.7×10^{-2}
Mean (weight)	176.25	132.5
Variance (weight)	1.22×10^2	5.5×10^2
Mean (Foot)	11.25	7.5
Variance (Foot)	9.7×10^{-1}	1.67

Classify the Person

- ▶ Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.

Classify the Person

- ▶ Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- ▶
$$P(F|6ft, 130lbs, 8units) = \frac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$$

Classify the Person

- ▶ Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- ▶
$$P(F|6ft, 130lbs, 8units) = \frac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$$

Classify the Person

- ▶ Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- ▶
$$P(F|6ft, 130lbs, 8units) = \frac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$$
- ▶
$$P(130lbs|F) = \frac{1}{\sqrt{2\pi \times 550}} \times \exp \frac{-(132.5-130)^2}{2 \times 550} = .0167$$

Classify the Person

- ▶ Given height = 6ft, weight = 130 lbs, feet = 8 units, classify if it's male or female.
- ▶
$$P(F|6ft, 130lbs, 8units) = \frac{P(6ft|F)P(130lbs|F)P(8units|F)P(F)}{P(130lbs, 8units, 6ft)}$$
- ▶
$$P(130lbs|F) = \frac{1}{\sqrt{2\pi \times 550}} \times \exp \frac{-(132.5-130)^2}{2 \times 550} = .0167$$
- ▶ Finally, we get probability of female given data is greater than the probability of class being male given data.